Early detection of pine wilt disease in *Pinus tabuliformis* in North China using a field portable spectrometer and UAV-based hyperspectral imagery

Run Yu[1], Lili Ren[1,2] and Youqing Luo[1,2]*

Abstract

Background: Pine wilt disease (PWD) is a major ecological concern in China that has caused severe damage to millions of Chinese pines (*Pinus tabulaeformis*). To control the spread of PWD, it is necessary to develop an effective approach to detect its presence in the early stage of infection. One potential solution is the use of Unmanned Airborne Vehicle (UAV) based hyperspectral images (HIs). UAV-based HIs have high spatial and spectral resolution and can gather data rapidly, potentially enabling the effective monitoring of large forests. Despite this, few studies examine the feasibility of HI data use in assessing the stage and severity of PWD infection in Chinese pine.

Method: To fill this gap, we used a Random Forest (RF) algorithm to estimate the stage of PWD infection of trees sampled using UAV-based HI data and ground-based data (data directly collected from trees in the field). We compared relative accuracy of each of these data collection methods. We built our RF model using vegetation indices (VIs), red edge parameters (REPs), moisture indices (MIs), and their combination.

Results: We report several key results. For ground data, the model that combined all parameters (OA: 80.17%, Kappa: 0.73) performed better than VIs (OA: 75.21%, Kappa: 0.66), REPs (OA: 79.34%, Kappa: 0.67), and MIs (OA: 74.38%, Kappa: 0.65) in predicting the PWD stage of individual pine tree infection. REPs had the highest accuracy (OA: 80.33%, Kappa: 0.58) in distinguishing trees at the early stage of PWD from healthy trees. UAV-based HI data yielded similar results: the model combined VIs, REPs and MIs (OA: 74.38%, Kappa: 0.66) exhibited the highest accuracy in estimating the PWD stage of sampled trees, and REPs performed best in distinguishing healthy trees from trees at early stage of PWD (OA: 71.67%, Kappa: 0.40).

Conclusion: Overall, our results confirm the validity of using HI data to identify pine trees infected with PWD in its early stage, although its accuracy must be improved before widespread use is practical. We also show UAV-based data PWD classifications are less accurate but comparable to those of ground-based data. We believe that these results can be used to improve preventative measures in the control of PWD.

Keywords: Pine wilt disease, Remote sensing, Spectrometer, Hyperspectral imaging, Random forest, Classification

* Correspondence: youqingluo@126.com
[1]Key Laboratory for Forest Pest Control, College of Forestry, Beijing Forestry University, Beijing 100083, China
[2]Sino-French Joint Laboratory for Invasive Forest Pests in Eurasia, Beijing Forestry University—French National Research Institute for Agriculture, Food and Environment (INRAE), Beijing 100083, China

Background

The pine wood nematode (PWN; *Bursaphelenchus xylophilus*) is a hazardous invasive species that infests multiple species of pine (Vicente et al. 2012; Douda et al. 2015). Pine wilt disease (PWD), caused by the PWN, is widespread throughout East Asia (Mamiya 1988; Hyun et al. 2007; Ye 2019). Previously isolated to southern China, PWD is now found throughout the country, including Northeast China (Pan et al. 2019; Yu et al. 2019). In 2016, PWD first appeared in Dalian, Liaoning Province, then in May 2017, it happened in Dandong City, Fushun City, Benxi City and other places (National Forestry Administration 2018). In addition, *Monochamus saltuarius* was identified as a new vector of PWD in Liaoning Province of China (Yu et al. 2018). In the process of spreading northwards, PWD has infected and caused severe damage to the Chinese pine (*Pinus tabulaeformis*), Korean pine (*P. koraiensis*), and larch (*Larix* spp.) populations. This has resulted in significant economic losses and ecological damage to Chinese pine forests (e.g., Li et al. 2011; Lin 2015; Hui 2018).

To effectively control PWD, it is necessary to identify infected trees in the early stage of infection. This is a difficult task because most trees progress from initial infection to the serious infection stage within 5 weeks (Umebayashi et al. 2017). Consequently, current PWD management strategies emphasize the control of infected trees after the onset of an outbreak by means of fumigation, burning, and tree felling (Shin 2008; Kim et al. 2018). What is lacking is a methodology that monitors pine populations that can quickly and efficiently detect the early signs of PWD (Ma et al. 2011). In addition, many efforts have been made in early detection of PWD (Kim et al. 2018; Syifa et al. 2020; Tao et al. 2020), but not in Chinese pine. In this paper, we present a method aimed at detecting PWD in Chinese pine in early stage.

A major obstacle in the management of pines infected by PWD is that the forests they persist in are very large communities. This can make classical ground identification and sampling methods impractical. To solve this problem, recent studies have used remote sensing (RS) to examine the impact of PWD on the physiological and biochemical changes after infection (e.g., Shen et al. 2001; Li et al. 2004; Wang et al. 2007). Advancements in RS technology increasingly support the prediction efficiency by reducing inherent spatial and temporal constraints (Ahmed et al. 2020). Similarly, hyperspectral remote sensing (HRS) can obtain continuous spectral information of objects – this has been used to detect changes in the spectral characteristics of needles on infected-trees in the process of discoloration caused by PWN infection (Pan 2011; Kuai 2012).

Previous studies show the presence of PWD is significantly linearly correlated with water and chlorophyll content. Therefore, water and chlorophyll content could be used as indicators of PWD (Huang 2020). This is important because RS and HRS methods can be used to estimate water and chlorophyll content. For example, using a field portable spectrometer to measure the spectral characteristics of *P. thunbergii* and *P. massoniana* at different stages of PWN infection, Xu et al. (2011) found the reflectance spectrum curve in the mid-infrared band may indicate the early stage of PWD with the analysis of the spectral characteristic parameters and changes in chlorophyll levels. Similarly, Xiang et al. (2018) used a field portable spectrometer, analyzing the relationship between spectral properties and chlorophyll, showing that the chlorophyll content of pine decreases with the stage of PWD (later, more severe stages are associated with lower chlorophyll content). In addition, the position of red edge, the wavelength of red edge, the height of green peak, and the depth of red band absorption all strongly correlate with chlorophyll content (e.g., Xiang et al. 2018). Correspondingly, the area surrounded by the first-order differential spectrum in the 490–530 nm range and that in the 680–760 nm range was found to be a significant hyperspectral feature indicating the occurrence of PWD (e.g., Huang et al. 2012). These studies all used a field portable spectrometer, which cannot be applied in a large-scale area.

Past studies used satellite imagery such as Landsat, IKONOS, Quick Bird, and GF-2 images to detect forest pest disease (e.g., Franklin et al. 2003; White et al. 2005; Hicke and Logan 2009; Zhan et al. 2020). However, due to limitations in spatial, temporal, spectral resolution as well as weather complications, satellite imagery cannot obtain real-time data (Santoso et al. 2016). Because of these limitations, the detection scale of forest pest disease has shifted to Unmanned Airborne Vehicle (UAV) remote sensing, which offers the advantages of low consumables and operating costs, high ground resolution data collection, and more precise accuracy (Tang et al. 2015). For example, Huang et al. (2018) used a fixed-wing unmanned aerial vehicle to monitor dead pine trees caused by PWD, successfully monitoring pine tree mortality with over 80% accuracy. Li et al. (2020) used UAVs to acquire remote sensing images of forest areas to assess the presence of PWD, successfully recognizing infection with 90.4% accuracy. Huang (2020) used UAV multispectral data to draw a conclusion that the first derivative of healthy and infected *P. thunbergii* changed markedly at 710 nm. Except RGB and multispectral camera, hyperspectral imagery was also applied in detecting forest pest diseases. Abdel-Rahman et al. (2014) used airborne hyperspectral data, random forest and support vector machines classifiers to distinguish amongst healthy, *Sirex noctilio* grey-attacked and lightning-damaged pine trees. Zhang et al. (2018) utilized the

ISIC-SPA-P-PLSR framework based on UAV-based hyperspectral image to identify the degree of damage trees caused by *Dendrolimus tabulaeformis*. Iordache et al. (2020) acquired airborne multispectral and hyperspectral data, and used Random Forest algorithms to compare the classification accuracies of the two datasets in detecting PWD, finding that both datasets performed well in identifying the infected, suspicious, and healthy trees. Importantly, however, in detecting the PWD, most studies focus on distinguishing between healthy and infected trees using RGB (red, green and blue bands) camera, multispectral data, and ground hyperspectral data, but UAV-based hyperspectral data were not widely studied. In addition, few studies emphasized the identification of trees in each stage of PWD infection in Chinese pine, which we focus on in this study. In our study, we systematically divided the infection stage into four stages, making the detection more accurate. Because high spatial and spectral resolutions, and feasibility of large-scale area application are needed to distinguish the subtle difference between healthy and the early stage of infected trees, we consider UAV-based hyperspectral imagery.

Spectral indices, such as Vegetation indices (VIs), red edge parameters (REPs), and moisture indices (MIs), can reflect the infection condition of PWD (e.g., Kim et al. 2018; Huang 2020). VIs is a combination of different remote sensing spectral bands, which can be regarded as a sign of relative abundance and activity of green vegetation (Jones and Vaughan 2010). Over the past years, VIs had been widely applied to extract sensitive estimates of plant biochemical characteristics (e.g., He et al. 2015; De Klerk and Buchanan, 2017), such as the normalized difference vegetation index (NDVI) that decreases with increasing tree PWD stage severity (Kim et al. 2018), and the presence of PWD can be detected by calculating the VIs based on ground, aerial, and satellite data (e.g., White et al. 2007; Pan et al. 2014; Jung and Park 2019; Iordache et al. 2020). The REPs are derived from Red edge (680–780 nm), which is the most obvious feature of plant spectral curve. As an indicator of plant stress and often used to study the growth and health of plants (Boochs et al. 1990; Dawson and Curran 1998), it also had been well studied in detecting the PWD (Du et al. 2009; Huang et al. 2012). Additionally, pine trees killed by PWD by blocking transmission of water (Yang 2002), and the water content of pine needles decreased with increasing PWD infection severity (Chen, 2005). Thus, changes in MIs (also derived from radiometric data) can be used to detect the presence of trees infected with PWD (Xu et al. 2012; Song et al. 2018).

Although spectral indices were widely used to detect the PWD, there is no study that yet provides good parameters to predict each stage (healthy, early, middle,

and serious stages) of PWD infection in Chinese pine trees. Additionally, analyses of PWD simultaneously considering ground and UAV-based hyperspectral data have not been widely conducted.

Therefore, to fill this gap, in this study, our objective is to explore the capacity of ground and airborne hyperspectral data using VIs, REPs, and MIs to classify the stage of PWD infection in Chinese pine at the tree level. Furthermore, we also aim to provide a useful and fairly accurate method of distinguishing between trees in the early stage of PWD infection from healthy trees.
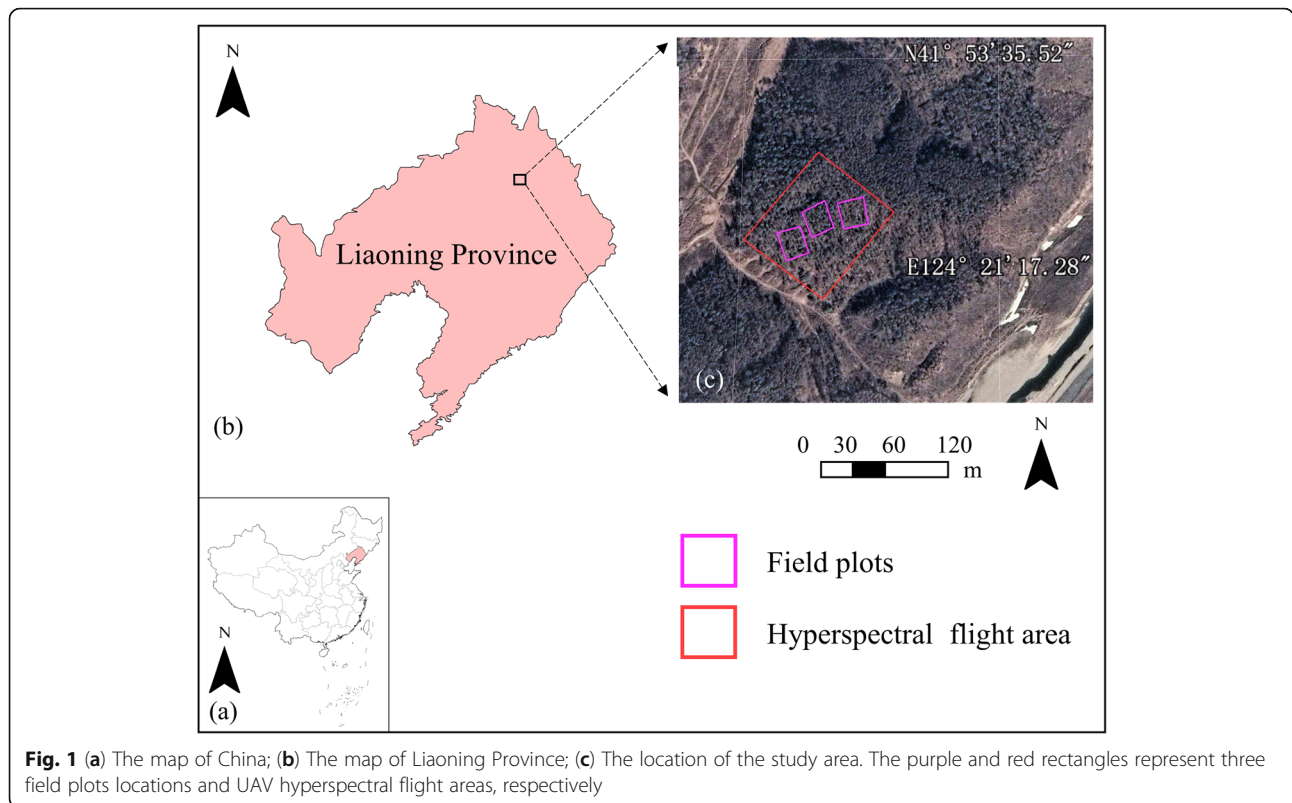
Materials and methods
Study area and ground survey
We conducted our study in Cangshi Village, located in Fushun County, Liaoning Province, in northeastern China (124°21′–124°24′ E, 41°53′–41°57′ N; Fig. 1). In the study area, the species of plantation forests is dominated by Chinese pine (*P. tabulaeformis*), and the age of them is approximately 40–50 years. The total area of forest cover in Fushun County is approximately 12.43×10^4 ha, of which *P. tabulaeformis* makes up > 30%. In addition, the broadleaf tree species and understory vegetation in the study site mainly include, *Quercus acutissima*, *Quercus mongolica*, grass, et al. The area is situated in the Middle Temperate Zone. It has a continental monsoon climate and experiences approximately 804.2 mm of precipitation per year. The mean annual air temperature is approximately 6.6 °C.

According to local Forestry Administration records, PWD has resulted in the death of tens of thousands of pine trees since the onset of outbreaks in 2016 in Liaoning Province (National Forestry Administration 2018).

Field measurements were conducted in 12–18 August 2019. We established three 30 m × 30 m plots located northeast of Cangshi Village (Fig. 1). The coordinates of the plot boundary and the location of each tree were recorded using a handheld differential global positioning system (DGPS, Version S760) with sub-meter accuracy. In each plot, we recorded tree growth state including tree height (H), diameter at breast height (DBH), crown diameter (CD), and PWD infection stage. In addition, we measured biochemical parameters: the leaf chlorophyll content (Cab) and water content (WC) of each tree. Cab was derived by averaging the Cab of needles from four different directions using a calibrated CCM-300 Chlorophyll Content Meter. The Cab of seriously damaged trees was 0 measured by the CCM-300. Meanwhile, WC of each tree was determined by the fresh weight (FW) minus the dry weight (DW) divided by the FW: (WC = (FW − DW)/FW × 100%). Finally, a total of 218 pine trees (healthy: 76; early stage: 54; middle stage: 47; serious stage: 41) were measured. Summary statistics of three plots are given in Table 1.

Yu *et al. Forest Ecosystems*   (2021) 8:44

Page 4 of 19

Fig. 1 (**a**) The map of China; (**b**) The map of Liaoning Province; (**c**) The location of the study area. The purple and red rectangles represent three field plots locations and UAV hyperspectral flight areas, respectively

Additionally, we randomly selected 20 discolored pine trees as samples from each plot, and took them back to the laboratory for testing by Behrman funnel method. The result showed that they all carried pine wood nematode.

Infected stage division

On the basis of previous studies (Xu et al. 2011; Santos and de Vasconcelos, 2012), we combined needle, ground tree, and UAV images to categorize PWD infection into four stages: (1) Healthy, (2) Early stage, (3) Middle stage, and (4) Serious stage (Fig. 2). Stages were defined by color of needles, growth vigor, and resin secretion (Table 2). We had four people classified each tree, and took the majority's opinion as final results to reduce subjective errors. Finally, we used the following definitions: "Healthy" trees

were defined as having dark green needles, normal resin secretion, and vigorous growth. "Early stage" trees were defined by slightly yellowed needles, with decreased resin secretion and grow rates. "Middle stage" trees were defined by yellow-brown needles, wilt, and weak growth. Dry trees with reddish-brown needles were defined as the "Serious stage".

Remote sensing data acquisition and preprocessing

Ground spectrum acquisition

From the ground (physically measuring trees in the field), we measured the spectrum of sampled trees using ASD Field Spec 4 Hi-Res NG (Analytical Spectral Devices, Boulder, CO, USA). The spectral range is 350–2500 nm and the spectral resolution is 3 nm in the 350–1000 nm wavelength range and 6 nm in the 1001–2500 nm

Table 1 Statistics of three plots variables (tree numbers = 218)

	Mean	Standard deviation	Maximum	Minimum	Range
H (m)	12.66	1.12	14.50	9.60	4.90
DBH (cm)	21.79	5.40	34.06	7.83	26.23
CD (m)	2.13	1.22	7.80	0.80	7.00
Cab (g·m^{-2})	299.30	165.93	560.00	0.00	560.00
WC (%)	42.13	14.45	68.32	17.25	51.07

H Tree height, *DBH* Diameter at breast height, *CD* Crown diameter, *Cab* Leaf chlorophyll content, *WC* Water content
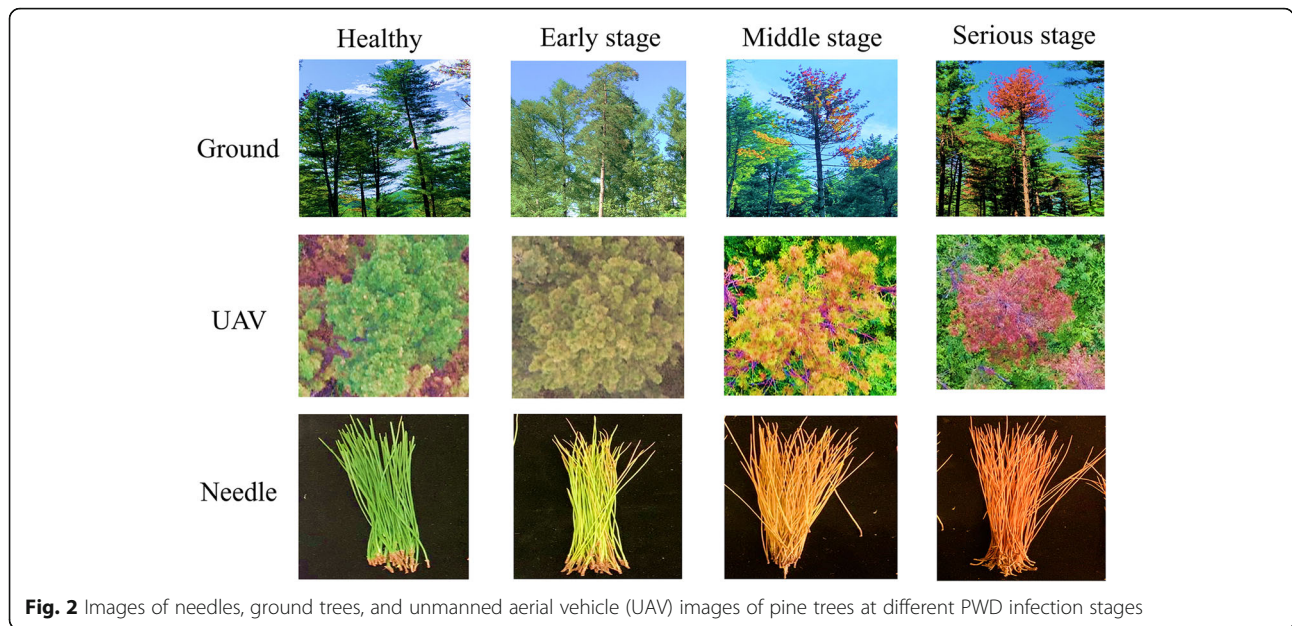
Fig. 2 Images of needles, ground trees, and unmanned aerial vehicle (UAV) images of pine trees at different PWD infection stages

wavelength range. We selected and measured branches roughly representative of the average spectrum of each tree. The selected branches were cut from the east, south, west, and north directions from the upper, middle, and lower layers (Zhang et al. 2018). We calculated the spectrum of each sampled tree by averaging the spectrum of the selected branches. The ground spectrums were gathered from 10:00 to 14:00 every day, from August 12 to August 17. We obtained the ground spectrum for comparison with UAV-based data and auxiliary radiometric correction.

UAV-based hyperspectral imagery
Hyperspectral Imagery (HI) data were obtained by using a DJI Matrice 600 UAV (DJI, Shenzhen, China) equipped a Pika L hyperspectral camera (Resonon, USA). The main parameters of the Pika L are listed in Table 3. GNSS (Global Navigation Satellite System) and IMU (Inertial Measurement Unit) modules are integrated into UAV, and its horizontal and vertical position errors are approximately 2.0 and 5.0 m, respectively, with an orientation precision of approximately 1 degree. The overall UAV-based system is shown in Fig. 3.

UAV-based hyperspectral data acquisition was carried out in the test areas of Cangshi Village from 12:00–12:

30, on 18 August 2019. The weather was sunny during the flight. Standard white board and white tarp were placed on the ground within the flying area. The flying height was set at 120 m, the image forward and side overlaps were set to 50%, and the flight speed is 2 m·s^{-1}. The imagery consisted of 281 spectral channels (spectral resolution of 2.1 nm) from visible to near infrared (NIR) regions (400–1000 nm). Reflectance correction and radiometric calibration were performed using 3 m^2 carpet reference (standard white board) and the Spectronon software. Image geometric corrections were performed using 4 ground control points (GCPs). The positions of GCPs were recorded by a DGPS device with sub-meter accuracy. The ground resolution of HI was produced to be 0.4 m.

Tree crowns extraction from hyperspectral imagery
We conducted tree crown segmentations from HI by combining the object-based segmentation method with manually drawing ROIs (regions of interest). First, by use of ENVI 5.3, we used the object-based segmentation method on the HIs using combined spectral and texture features to separate trees crowns from the grass background and shadows (e.g., Yuan et al. 2013). The object-based segmentation method successfully separated tree

Table 2 Classification of infected stages

Age of the stand	Classification standard	Stages of infected trees			
		Healthy	Early stage	Middle stage	Serious stage
40–50 years	Color of needles; Growth vigor; Resin secretion	Trees grow vigorously with dark green needles; resin secretes normally	Needles begin to turn yellow; resin secretion decreases and growth slows down	Most needles turn yellow brown and wilt, and the growth is obviously weak	Trees are dry, and needles are all reddish brown, but do not fall off

Yu *et al. Forest Ecosystems*   (2021) 8:44

Page 6 of 19

Table 3 Main parameters of the Pika L imaging spectrometer (provided by manufacturer)

Parameters	Values	Parameters	Values
Weight	0.6 kg	Sampling interval	1.07 nm
Digitization	12 bits	Spectral resolution	2.10 nm
Wavelength range	400–1000 nm	Spectral channels	281

crowns from the grass background and shadow components. However, it was difficult to separate overlapping crowns. Second, based on the result of object-based segmentation, we drew the ROIs manually. We determined the location of every individual sampled trees by use of the DGPS information. The ROIs of each tree were shaped by manually drawing the crown range on the RGB image. Then, the ROIs were added to the preprocessed HIs, and the spectrum of an individual tree was calculated by averaging the reflectance of the corresponding ROI extracted by ENVI 5.3. The average spectrum information of each ROI was used in the subsequent analysis (Fig. 4). Finally, the shadow components and overlapped crowns were discarded. Overall, 121 trees (healthy: 39; early stage: 27; middle stage: 29; serious stage: 26) were segmented from HI hyperspectral imagery.

Features extraction

In order to eliminate instrument errors and noises, while maintaining the original spectral characteristics, a Savitzky-Golay filter with 7 points (we tested 3–15 points and finally chose 7 points) was used to smooth the original spectrums of ground and UAV-based hyperspectral data (Mullen 2016). Based on previous research, we calculated 37 spectral variables including 12 VIs (Table 4), 20 REPs showed in Table 5 (Horler et al. 1980; Curran et al. 1990; Yao et al. 2009; Liu et al. 2010), and 5 MIs (Table 6) from spectral data.

Feature selection and prediction model for cab and WC

According to previous study (De Klerk and Buchanan, 2017; Kim et al. 2018; Lin et al. 2019), trees health was highly correlated with biochemical properties (e.g., Cab), which also can be precisely estimated using fitting models based on spectral indices (Inoue et al. 2012; Schlemmer et al. 2013; Xie et al. 2014). In this study, we firstly calculated the Pearson's correlation coefficient between a number of spectral indices (features in Tables 4, 5 and 6) and Cab and WC, PWD infection stages of each tree, respectively. In addition, before these variables were selected for constructing regression and classification model to predict Cab, WC, and PWD infection stages, we used a stepwise regression method to test the multicollinearity between features, eliminating redundant variables.

Finally, we selected 5 VIs (NDVI, NDVI [810, 680], NDVI [560, 680], RVI, PRI), 5 REPs (λo, Sg, Kg, GH, RD), and 5 moisture indices (MSI, WI1, WI2, NDWI, NSII) based on ground spectrum (350–2500 nm). Based on UAV hyperspectral data (400–1000 nm), we selected 5 VIs (PSI, RVSI1, NDVI, NDVI [810, 450], RVI), 5 REPs (dλb, SDr, SDb, SDr-SDb, RD) and 2 MIs with (WI1, WI2).

We estimated the Cab and WC using a RF (Breiman 2001) regression using a bagging method based on the CART regression tree model. In the regression application, each tree was built by choosing a random sample and a random set of variables from the training dataset by a deterministic algorithm (Mutanga et al. 2012). All 121 samples were used for model training, and we then used a 10-fold cross-validation method (Waske et al. 2009) to assess model accuracy. The process of regression was conducted using the R package "randomForest". The coefficient of determination (R^2), RMSE (Root Mean Square Error), and RRMSE (Relative RMSE) between measured and estimated values were used to compare different indices in predicting the accuracy of Cab and WC. After selecting the variables which performed best in predicting the Cab and WC, we used the Cab or
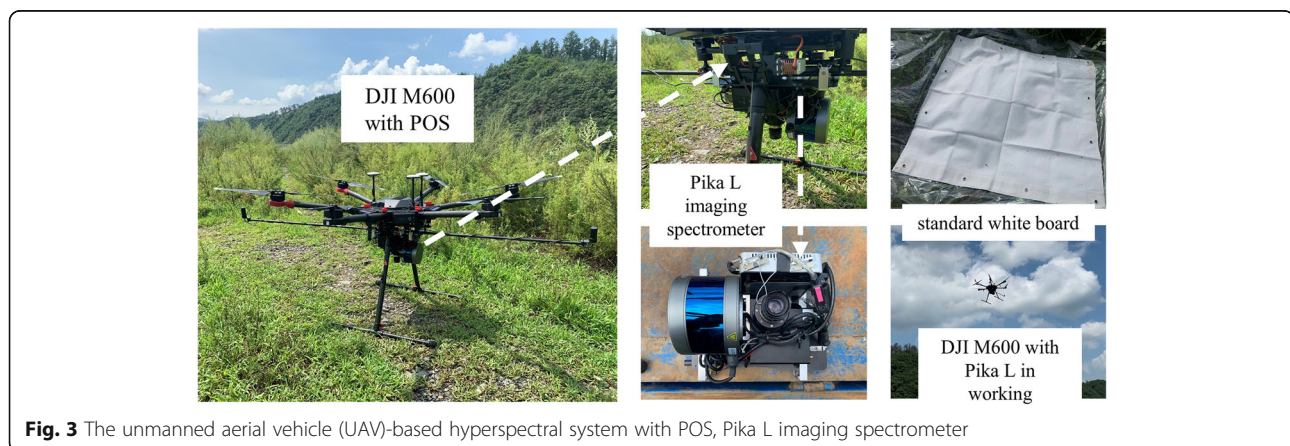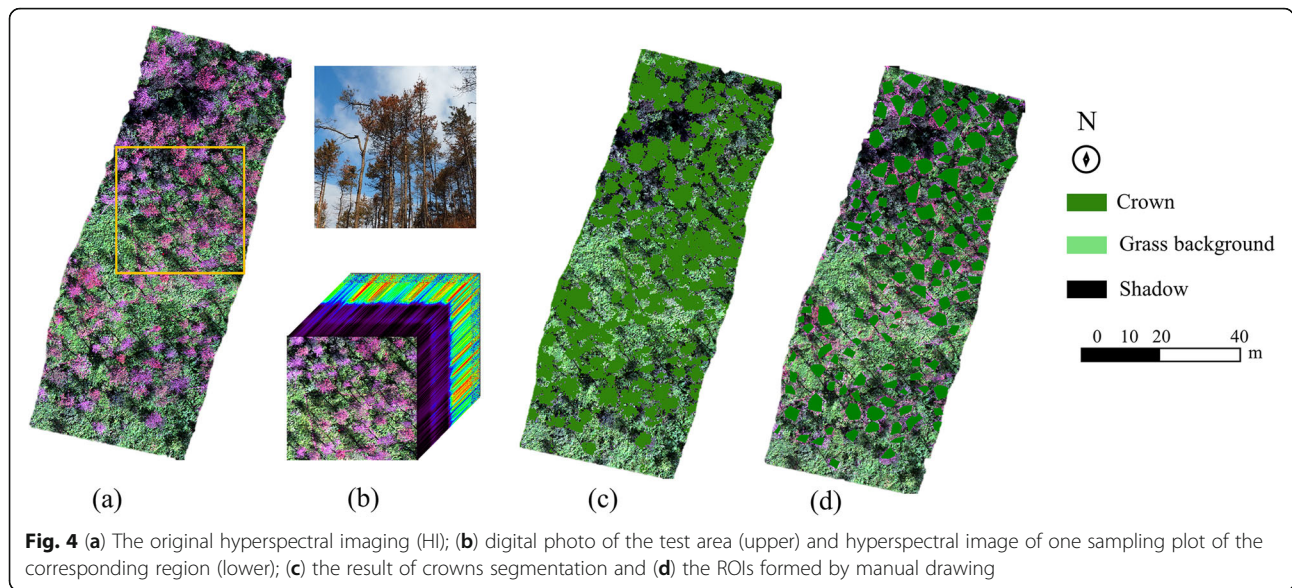

Fig. 3 The unmanned aerial vehicle (UAV)-based hyperspectral system with POS, Pika L imaging spectrometer

Fig. 4 (**a**) The original hyperspectral imaging (HI); (**b**) digital photo of the test area (upper) and hyperspectral image of one sampling plot of the corresponding region (lower); (**c**) the result of crowns segmentation and (**d**) the ROIs formed by manual drawing

WC estimated by the optimum variables to classify the PWD infection stages directly.

Classification based on VIs, REPs, and MIs

We then used the selected VIs, REPs, and MIs correlated with Cab and WC to classify trees based on PWD infection. We used a RF classification model to assess the infection stage of sampled trees. In a RF algorithm, the variable importance is a metric of how much the "out-of-bag" (OOB) error of estimate increases due to the removal of a single variable from the data (Prasad et al. 2006; Verikas et al. 2011). The mean decrease accuracy (MDA) index of each variable is obtained when calculating the OOB error:

the higher the MDA value of a variable is, the more important it is (e.g., Liu et al. 2017; Shi et al. 2018).

The selected VIs, REPs, MIs and combining all variables were separately input into RF classification model, and the MDA of all selected variables were determined. All 121 samples were used for model training. We then used a 10-fold cross-validation method to estimate model accuracy. The process of classification was carried out using the R package "randomForest". The overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA), and Kappa coefficient resulting from confusion matrices (Congalton 1991) were used to evaluate classification accuracy. Kappa coefficient is a popular statistic

Table 4 Vegetation indices extracted from spectral data

Variables	Description	Formula	Reference
NDVI	Normalized difference vegetation index	$\frac{\text{SUM } R(760:900)/141 - \text{SUM } R(630:900)/271}{\text{SUM } R(760:900) + \text{SUM } R(630:900)/271}$	Richardson and Wiegand (1977)
NDVI (810,450)	Normalized difference vegetation index	$\text{NDVI}(810,450) = \frac{(R_{810} - R_{450})}{(R_{810} + R_{450})}$	Richardson and Wiegand (1977)
NDVI (810,680)	Normalized difference vegetation index	$\text{NDVI}(810,680) = \frac{(R_{810} - R_{680})}{(R_{810} + R_{680})}$	Richardson and Wiegand (1977)
NDVI (560,680)	Normalized difference vegetation index	$\text{NDVI}(560,680) = \frac{(R_{560} - R_{680})}{(R_{560} + R_{680})}$	Richardson and Wiegand (1977)
RVI	Ratio vegetation index	$\frac{\text{SUM } R(760:900)/141}{\text{SUM } R(630:900)/271}$	Wu and Niu (2008)
DVI	Deferent vegetable index	$\text{DVI} = \text{SUM } R(760:900)/141 - \text{SUM } R(630:900)/271$	Chen (1996)
PRI	Photochemical reflectance Index	$\text{PRI} = \frac{R_{570} - R_{531}}{R_{570} + R_{531}}$	Carter and Miller (1994)
MSR	Modified simple ratio	$\text{MSR} = \frac{\frac{R_{800}}{R_{670}} - 1}{\text{sqrt}(\frac{R_{800}}{R_{670}} + 1)}$	Blackburn (1998)
PSI	Plant stress index	$\text{PSI} = \frac{R_{695}}{R_{760}}$	Hunt and Rock (1989)
RVSI1	Ratio vegetation stress index	$\text{PVSI1} = \frac{R_{600}}{R_{760}}$	Hunt and Rock (1989)
RVSI2	Ratio vegetation stress index	$\text{PVSI2} = \frac{R_{710}}{R_{760}}$	Hunt and Rock (1989)
PSSR	Pigment specific simple ratio	$\text{PSSR} = \frac{R_{800}}{R_{635}}$	Penuelas et al. (1997)

Table 5 Red Edge parameters extracted from hyperspectral data

Type	Parameters	Description	Type	Parameters	Description
Based on original spectrum	Rg	The maximum reflectance in the wavelength range of 510–560 nm	Based on first derivative	dλb	The first order differential value corresponding to λb
	Ro	The minimum reflectance in the wavelength range of 640–680 nm		dλr	The first order differential value corresponding to λr
	λg	Wavelength position of Rg		SDb	The area surrounded by the first-order differential spectrum in the range of 490–530 nm
	λo	Wavelength position of Ro		SDr	The area surrounded by the first-order differential spectrum in the range of 680–760 nm
	Sg	Skewness of reflectance in the wavelength range of 510–560 nm		SDr–SDb	SDr–SDb
	Kg	Kurtosis of reflectance in the wavelength range of 510–560 nm	Others	Lwidth	Width at half depth of red band absorption
	Sr	Skewness of reflectance in the range of 680–760 nm		Depth672	Absorption depth at 672 nm
	Kr	Kurtosis of reflectance in the wavelength range of 680–760 nm		Depth560	Absorption depth at 560 nm
Based on first derivative	λb	Wavelength position of the maximum first derivative of reflectance between 490 and 530 nm		GH	Height of green peak
	λr	Wavelength position of the maximum first derivative of reflectance between 680 and 760 nm		RD	Depth of red band absorption

for measuring agreement (Meddens et al. 2011). A Kappa value from < 0.4 indicates a "poor" agreement, Kappa 0.4–0.8 is defined as having moderate agreement, and Kappa > 0.80 indicates a "strong" agreement.

Using the overall and individual accuracies for all four PWD infection stages, we examined the paired accuracies of Healthy, Early stage, Middle stage, and Serious stage pine trees to examine the feasibility of discriminating between different stages.

Results
Estimation of cab and WC
Leaf Cab and WC decreased with the severity of PWD infection (Fig. 5). We estimated the Cab and WC of all 121 sampled trees using the RF regression model with the three input parameters (VIs, REPs, and MIs) separately

input. We examined the performance of Cab and WC estimation of the input parameters using both ground spectrum data and UAV-based spectral data (Figs. 6 and 7). Cab estimation accuracy was slightly greater when using REPs than using VIs for both ground data (REPs: $R^2 = 0.78$, RMSE = 82.34 g·m^{-2}, RRMSE = 27.44%; VIs: $R^2 = 0.74$, RMSE = 89.80 g·m^{-2}, RRMSE = 29.92%) and UAV-based data (REPs: $R^2 = 0.75$, RMSE = 87.34 g·m^{-2}, RRMSE = 29.11%; VIs: $R^2 = 0.72$, RMSE = 94.11 g·m^{-2}, RRMSE = 31.36%). For WC predictions in which MIs were used as input parameters, the predictions from ground data were considerably more accurate than UAV-based data. The results summarized in Table 7.

It showed that the model tended to overestimate Cab below 200 g·m^{-2} and underestimate Cab above 300 g·m^{-2} (Fig. 6a and b; Fig. 7a and b), the RF regression

Table 6 Moisture indices extracted from spectral data

| Variables	Description	Formula	Reference
MSI	Moisture stress index	$MSI = \frac{R_{1600}}{R_{820}}$	Gao (1996)
WI1	Water index	$WI1 = \frac{R_{970}}{R_{900}}$	Hardisky et al. (1983)
WI2	Water index	$WI2 = \frac{R_{950}}{R_{900}}$	Hardisky et al. (1983)
NDWI	Normalized difference water index	$NDWI = \frac{(R_{860} - R_{1240})}{(R_{860} + R_{1600})}$	Prasad et al. (2006)
NDII	Normalized difference infrared index	$NDII = \frac{(R_{860} - R_{1600})}{(R_{860} + R_{1600})}$	Verikas et al. (2011)
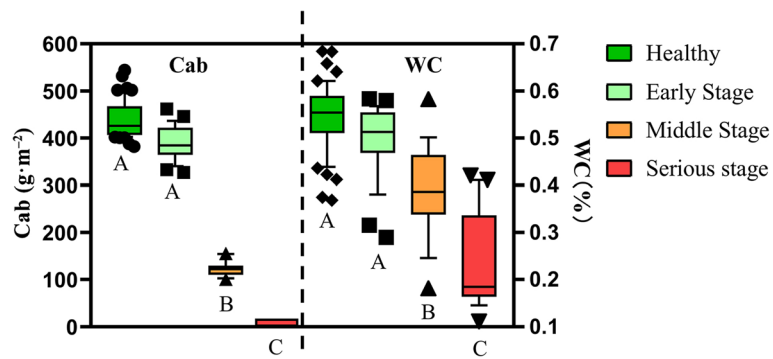
Fig. 5 The difference of chlorophyll content (Cab) and water content (WC) of all samples at different infected stages

model provided unsatisfactory predictions for Cab and WC in pine trees when VIs, REPs, and MIs were taken as input parameters.

In addition, Cab estimated by REPs derived from ground data (the optimum variables) were chosen to assess the PWD infection stages directly. Finally, the results showed that using Cab estimated by RF based on the optimum variables did not perform well in classifying the PWD infection stages (OA = 47.11%, Kappa = 0.29; Table 8). It means that estimated Cab cannot be directly used to accurately the PWD infection stages.

Feature analysis

The spectral reflectance of trees declined as a function of PWD stage severity (Fig. 8). The difference of spectral reflectance was obvious near the green peak (500–600 nm), red edge (680–760 nm), and NIR (750–950 nm; Fig. 9). VIs, REPs, and MIs exhibited differing responses to the severity of infection. While some variables such as NDVI (810, 680), Kg, and NDWI decreased with the increasing infection stage, others (e.g. MSI, PRI and Sg) significantly increased with the increasing of the infection stage (Fig. 10). Therefore, almost all the selected variables exhibited statistically significant responses to PWD severity, indicating their potential for detecting the

stage of PWD. Generally, the spectral variables were sensitive to changes in biochemical characteristic.

Comparisons of classifications using different variables from ground and UAV-based data

The MDA index for the ground data and UAV-based data strongly differed among variables. Importance rankings indicated REPs to be more important than most VIs and MIs (Fig. 11). The most important variables were REPs, and VIs were generally more important than MIs. GH was the most important variable for ground data and SDR was the most important variable for UAV-based data.

OA (overall accuracy) assessment using the 10-fold cross-validation method indicated that REPs performed best. For ground data REPs yielded an OA of 79.34%, VIs 75.21%, and MIs 74.38%. Combined all variables, it yielded an accuracy of 80.17 (Tables 9 and 10). UAV-based data provided less accurate results for all variables: 72.73% for REPs, 70.25% for VIs, 63.64% for MIs, and 74.38% for combined all variables (Tables 9 and 10). Kappa values yielded similar qualitative results for both ground data and UAV-based data. For ground data, Kappa was calculated to be 0.67 for REPs, 0.66 for VIs, and 0.65 for MIs. For combining all variables, Kappa improved to 0.73. For UAV-based data, the values of Kappa
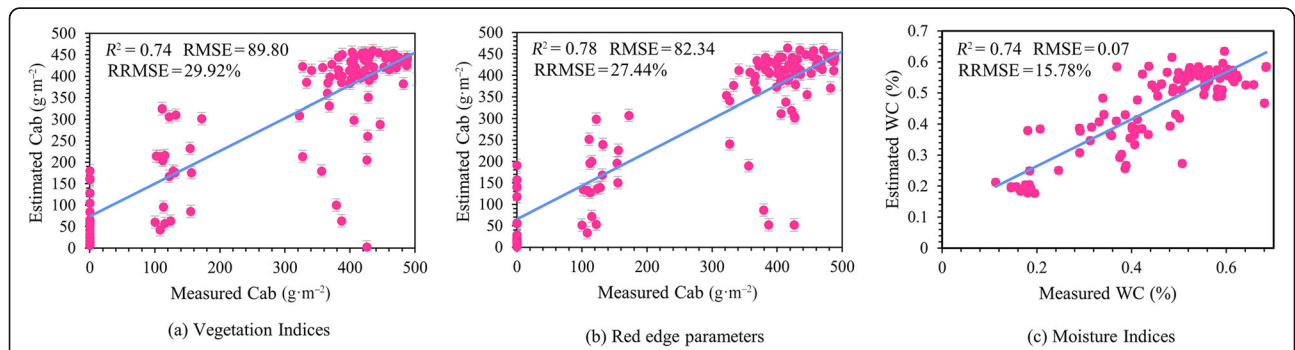

Fig. 6 Measured vs. estimated chlorophyll content (Cab) and water content (WC) based on ground spectrum using different input parameters: (**a**) Vegetation Indices; (**b**) Red edge parameters, and (**c**) Moisture Indices

Fig. 7 Measured vs. estimated chlorophyll content (Cab) and water content (WC) using UAV-based spectrum with different input parameters: (**a**) Vegetation Indices; (**b**) Red edge parameters, and (**c**) Moisture Indices
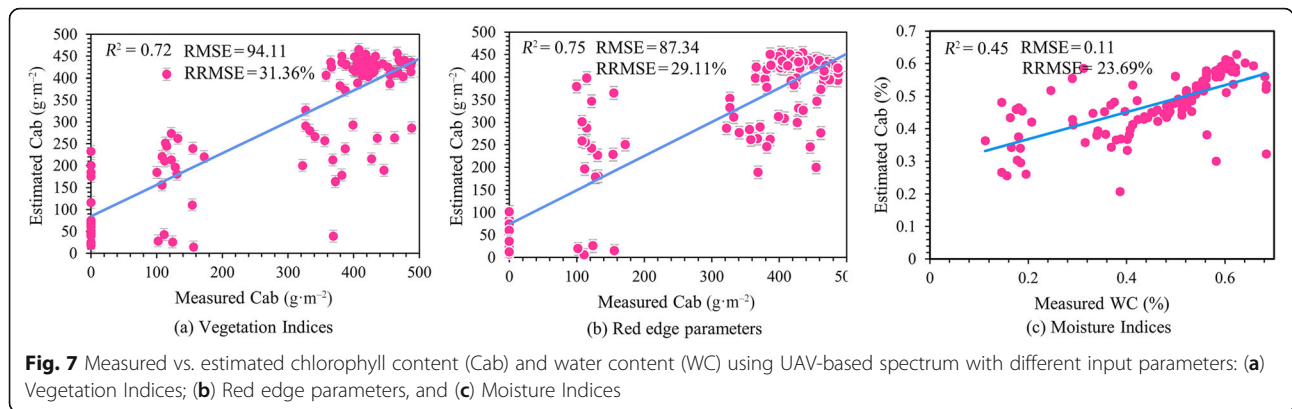
for REPs, VIs, and MIs were 0.63, 0.60, and 0.51, respectively. For combining all variables, Kappa again improved (0.66). Therefore, for each data type (ground data and UAV-based data), REPs yielded the most accurate results, followed by VIs and MIs respectively. Additionally, ground data provided more accurate results than UAV-based data in all cases.

PA (producer's accuracy) values were high for the middle and serious stage of infection regardless of the data source and the parameters used. UA (user's accuracy) was relatively high for middle and serious stage of infection, while healthy and early stages had lower UAs (Table 11).

Pairwise comparisons of healthy, early stage, middle stage, and serious stage indicated the OAs of all stage pairs to be considerably greater than 80% in most cases (Figs. 12 and 13). Lower accuracies resulted when healthy pine trees and early stage of infected pine trees were compared based on the VIs (75.41%), REPs (80.33%), MIs (70.97%), and combined all variables (79.03%) from ground data, as well as VIs (68.33%), REPs (71.67%), MIs (66.67%), and combined all variables (71.67%) from UAV-based data. High values of Kappa were obtained by most pairwise comparisons (Figs. 12 and 13), but not for comparisons between healthy pine trees and pine trees in the early stage of infection for both ground data (VIs: 0.55, REPs: 0.58, MIs: 0.39, combined all variables: 055) and UAV-based data (VIs: 0.35, REPs: 0.40, MIs: 0.31, combined all variables: 0.40). REPs

(OA: 80.33%, Kappa: 0.58) based on ground data performed best when healthy pine trees and early stage of infected pine trees were compared. REPs and combining all variables performed equally well in terms of OA (71.67%) and Kappa (0.40) for UAV-based data.

Discussion

In this paper, we employed VIs, REPs, MIs, and combining all variables, to examine the capacity of ground and UAV-based hyperspectral data in PWD infection stages estimation at individual tree level. The results reveal that combining all variables performed best and yielded a considerably accurate classification with OA of 80.17% for ground data and 74.38% for UAV-based data (Tables 9 and 10).

When we look at the capacity of identifying pine trees in the early infected stage of PWD, the REPs exhibited the best performance with OA of 80.33% and 71.67% from ground data and UAV-based data, respectively (Figs. 12 and 13).

Overall, it is understandable that: (1) the REPs are more responsive to stage changes of PWD infection than VIs and MIs, indicating that REPs may be more sensitive to the biochemical conditions; (2) UAV-based data performed considerable accuracy in monitoring the PWD stage at individual tree level, especially REPs, showing its good accuracy, which were slightly lower than ground data and can be applied in a large-scale forest area.

Table 7 The results of Cab and WC estimation using RF regression based on VIs, REPs, and MIs from ground and UAV hyperspectral data

Variable type	Cab and WC estimation					
	Ground			UAV		
	R^2	RMSE	RRMSE (%)	R^2	RMSE (%)	RRMSE (%)
VIs (estimating Cab)	0.74	89.80	29.92	0.72	94.11	31.36
REPs (estimating Cab)	0.78	82.34	27.44	0.75	87.34	29.11
MIs (estimating WC)	0.74	0.07	15.78	0.45	0.11	23.69

Table 8 Estimation of PWD infection stages using Cab estimated by RF based on the REPs derived from ground data

Stage	H	E	M	S	Total
H	25	11	0	0	36
E	12	4	6	0	22
M	1	12	14	12	39
S	1	0	9	14	24
Total	39	27	29	26	121
				OA (%)	47.11
				Kappa	0.29

Error analysis

Previous studies show hyperspectral data to be effective in examining forest health (e.g., Pontius et al. 2008; Näsi et al., 2015). However, we encountered several difficulties, obstacles, and sources of error in precisely estimating leaf Cab, WC, and the stage of PWD in pine trees. (1) The stage of PWD of each sampled pine tree was judged by visual observation. These measurements were fairly subjective and possibly inaccurate. (2) The acquisition of ground and UAV-based hyperspectral data are both easily affected by the weather, especially light. Because data were collected during light hours, this may have biased results. (3) The results of individual tree crown segmentation using UAV-based hyperspectral data were somewhat inaccurate. This increased the uncertainly of extracting tree hyperspectral features and, consequently, it was difficult to distinguish pine trees from understory trees and separate overlapping crowns from HIs using the image classification algorithm. Manually drawing and visual interpretation can reduce the interference of mixed pixels, but there was a problem that it cannot be efficiently applied when the sample size was large. Nevertheless, in the actual situation, we can hardly meet two requirements at the same time: obtaining pure pixels and those that completely cover the whole crown. (4) We collected Cab and WC data on 12–18 August 2019, while we acquired the UAV-based data on 18 August 2019. During the interval, the biochemical conditions may have changed. Because it only took 30–60 min for the drone to complete the data collection, but the artificial ground survey took at least 1 week. In this study, we cut each tree branch and then measured the spectrum, Cab, and WC of each tree. Therefore, the workload is relatively heavy, the ground survey cannot be synchronized with the drone data collection, and we can only keep the time as close as possible. (5) The results of our study may be affected by small sample size.

Possible application of UAV-based hyperspectral data in detecting PWD

Overall, the PWD infection stage classification of ground data was more accurate than that of UAV-based airborne hyperspectral data (Tables 9 and 10). There are several possible sources of this discrepancy. Firstly, ground data consisted of samples from the entire tree while the airborne data only measured canopy spectral data. Therefore, ground data samples may more accurately reflect the tree condition. Additionally, airborne data acquisition is easily affected by weather – this may have induced measurement errors. PA and UA of the four PWD infection classes using RF based on VIs, REPs, MIs, and combining all variables also suggest ground data performed better than airborne data (Table 11). However, when the RERs and combining all variables were used from UAV-based data, predictions were comparably accurate to those of ground data (Tables 9 and 10).

Importantly, the acquisition of airborne data is simple, convenient, and much faster than ground data acquisition. Therefore, there is a trade-off between the accuracy and efficiency of data acquisition: ground data acquisition is accurate but time consuming to obtain while UAV-based airborne data is less accurate but much
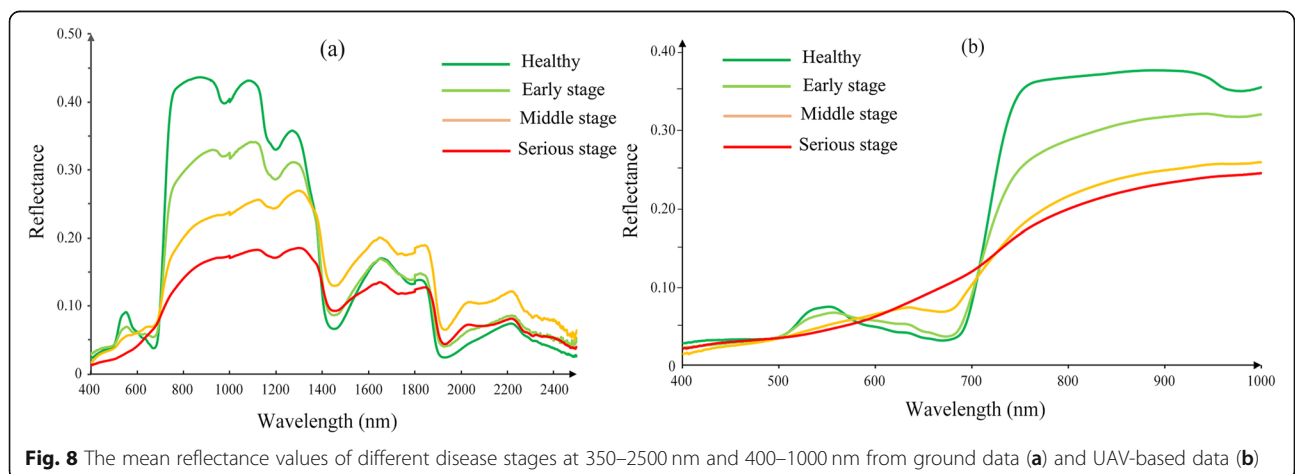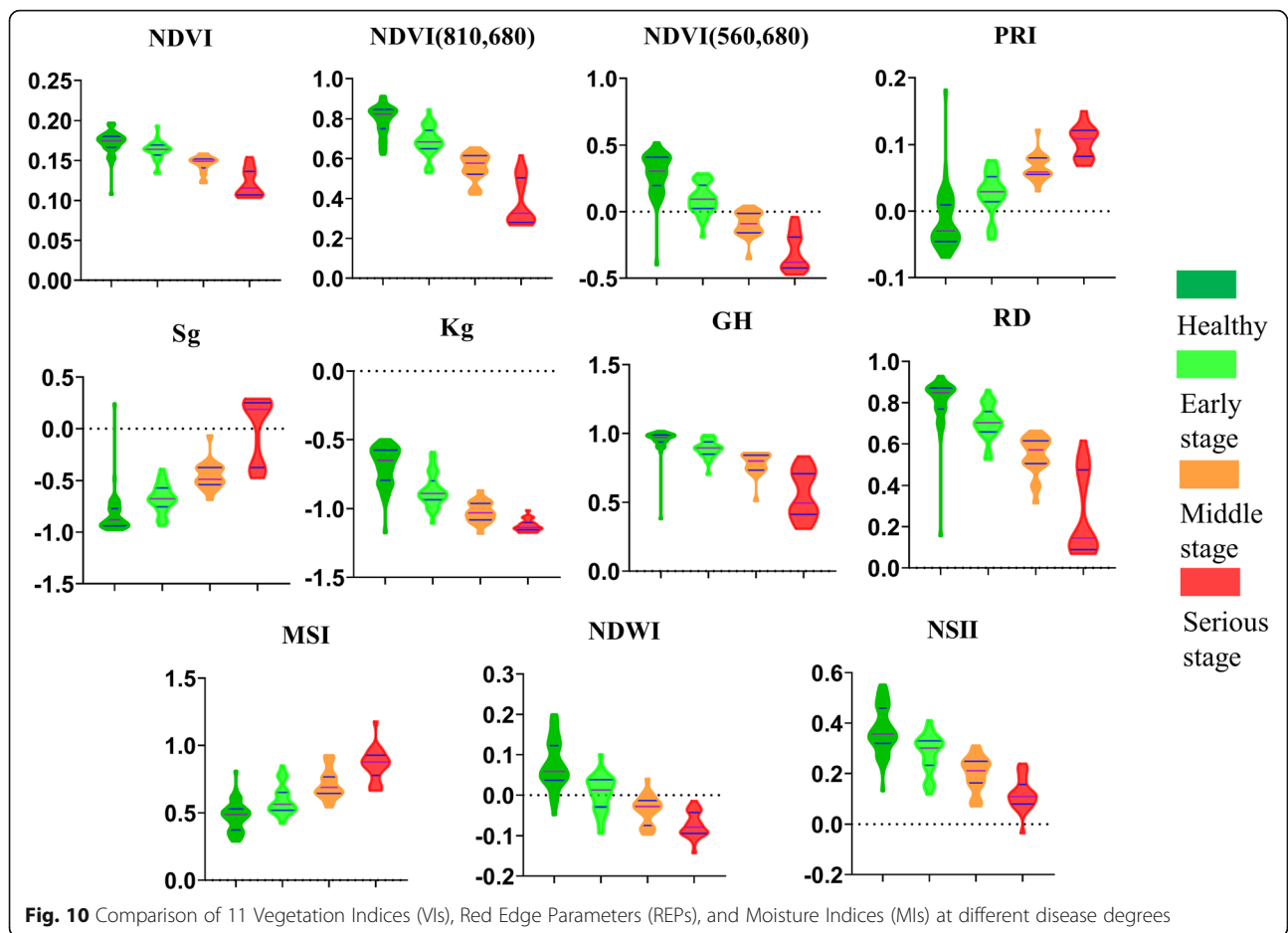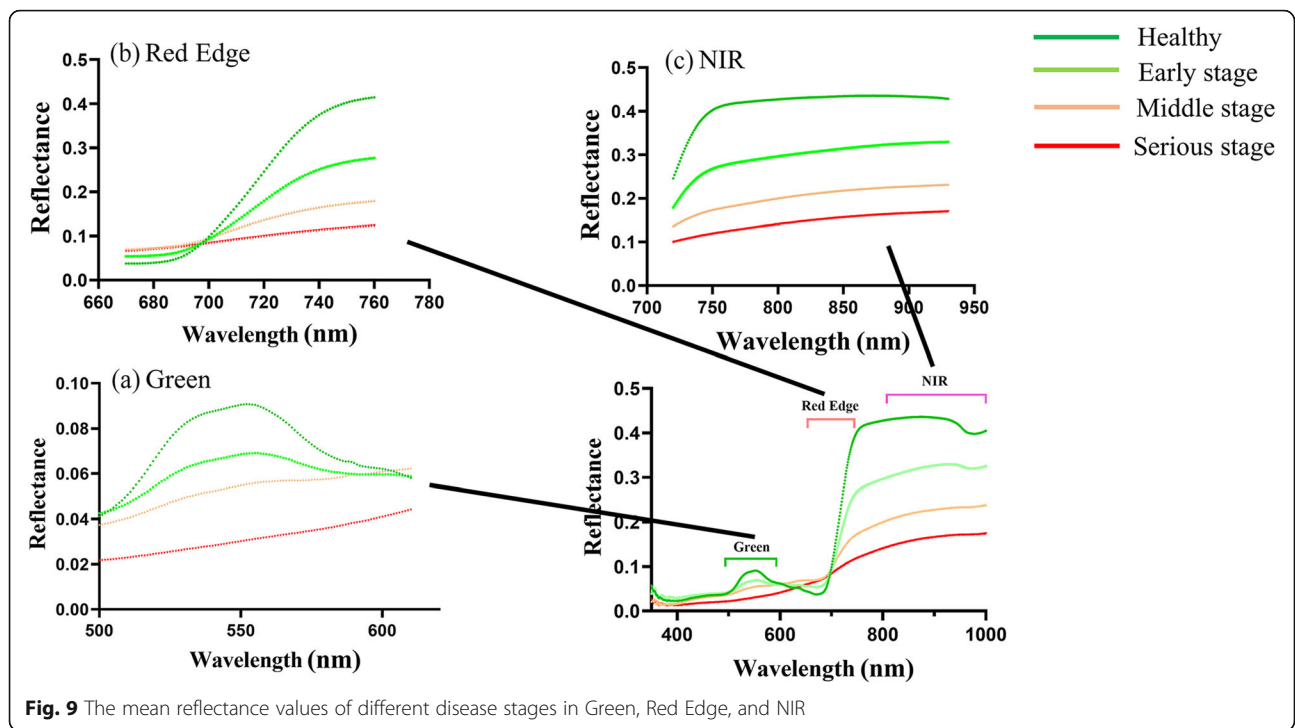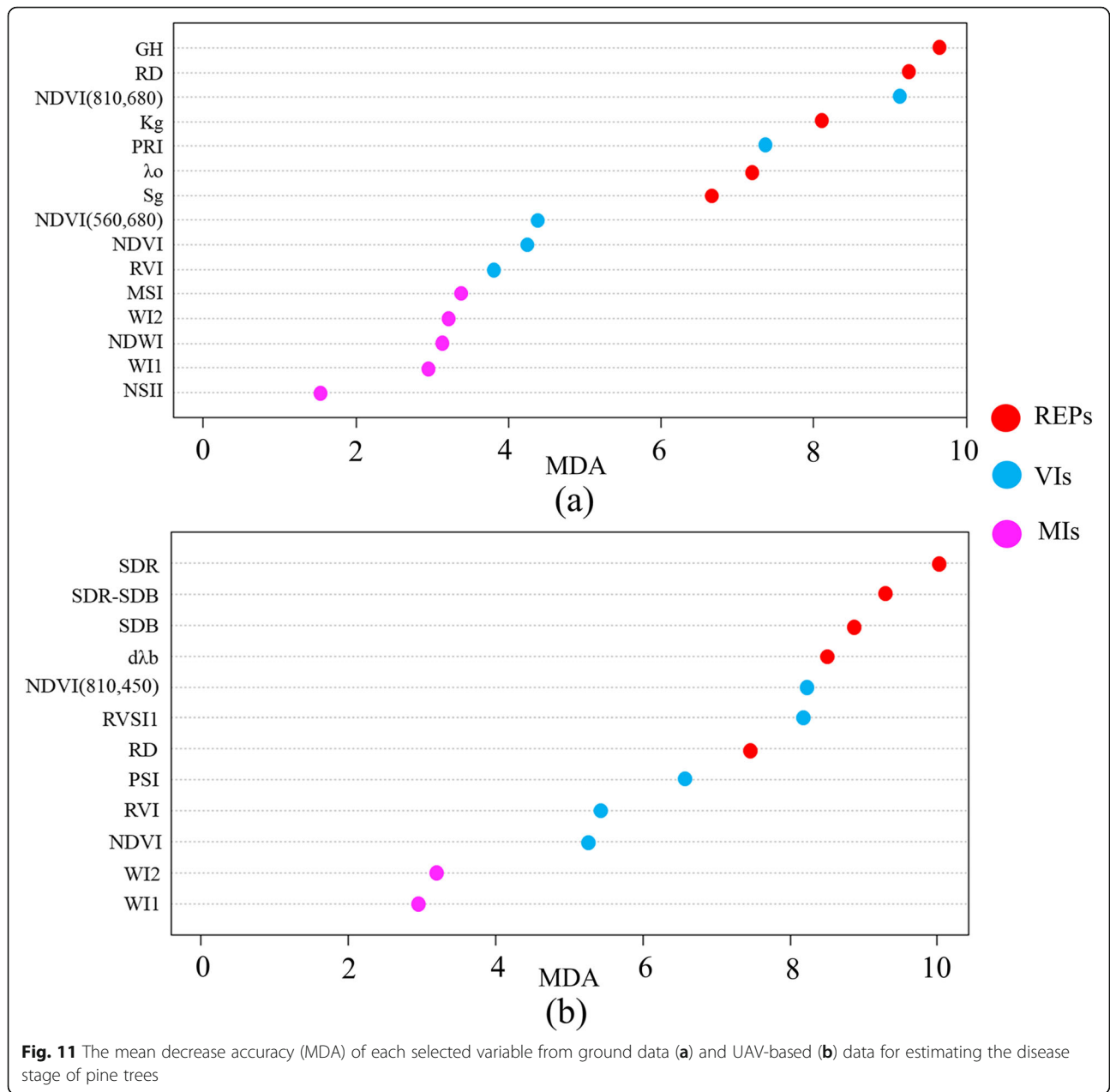

Fig. 8 The mean reflectance values of different disease stages at 350–2500 nm and 400–1000 nm from ground data (**a**) and UAV-based data (**b**)

Fig. 9 The mean reflectance values of different disease stages in Green, Red Edge, and NIR

Fig. 10 Comparison of 11 Vegetation Indices (VIs), Red Edge Parameters (REPs), and Moisture Indices (MIs) at different disease degrees

Fig. 11 The mean decrease accuracy (MDA) of each selected variable from ground data (**a**) and UAV-based (**b**) data for estimating the disease stage of pine trees

easier to obtain. Because PWD potentially affects trees in many large forest areas, ground data acquisition is not a feasible management strategy. UAV-based data provides only slightly less accurate classifications than ground-based data and is thus a more practical candidate for future large-scale forest management.

The potential of identifying trees in the early stage of PWD

Our results show that it is relatively simple to distinguish healthy trees and trees in early stage of PWD infection from trees in the middle and serious stage of PWD infection. This is because the biochemical

characteristics (e.g. leaf Cab) of healthy trees and trees at early stage of PWD are very different from those of trees in middle and serious stage (Fig. 5). In contrast, it is difficult to distinguish healthy trees from trees in early stage of PWD because the difference in their spectral responses cannot be detected easily. REPs performed relatively well in distinguishing trees in early stage of PWD infection from healthy trees (ground data OA: 80.33%, Kappa: 0.58; and airborne data OA: 71.67%, Kappa: 0.40); however, overall, UAV-based data yielded moderately low accuracy (Fig. 13). Therefore, in practical application, especially in a large-scale forest area, it is still a challenge to use UAV-based hyperspectral data to

Table 9 Classification confusion matrix of random forest (RF) classifier using vegetation indices, red edge parameters, moisture indices, and combined all variables based on ground spectral data

Stage	Vegetation Indices						Red Edge Parameters					
	H	E	M	S	Total	UA (%)	H	E	M	S	Total	UA (%)
H	29	7	0	0	36	80.56	32	7	0	0	39	82.05
E	8	17	3	0	28	60.71	5	17	1	0	23	73.91
M	1	2	22	3	28	78.57	1	2	24	3	30	80.00
S	1	1	4	23	29	79.31	1	1	4	23	29	79.31
Total	39	27	29	26	121		39	27	29	26	121	
PA (%)	74.36	62.96	75.86	88.46	OA (%)	75.21	82.05	62.96	82.76	88.46	OA (%)	79.34
					Kappa	0.66					Kappa	0.67
Stage	Moisture Indices						Combined					
	H	E	M	S	Total	UA (%)	H	E	M	S	Total	UA (%)
H	29	9	1	0	39	74.36	32	8	0	0	40	80.00
E	9	15	1	1	26	57.69	5	17	2	0	24	70.83
M	0	3	24	3	30	80.00	2	2	24	2	30	80.00
S	1	0	3	22	26	84.62	0	0	3	24	27	88.89
Total	39	27	29	26	121		39	27	29	26	121	
PA (%)	74.36	55.56	82.76	84.62	OA (%)	74.38	82.05	62.96	82.76	92.31	OA (%)	80.17
					Kappa	0.65					Kappa	0.73

precisely identify trees at early infected stage of PWD. In conclusion, the main focus of our next study is to improve the accuracy by some effective approaches (e.g., using multi-temporal UAV hyperspectral data).

Classification algorithms

Machine learning algorithms, such as Random forest (RF), support vector machine (SVM), Classification and Regression Tree (CART), have been widely conducted in classifying damaged trees by forest pest in previous studies (Abdel-Rahman et al. 2014; Iordache et al. 2020; Syifa et al. 2020; Zhan et al. 2020). In our study, RF algorithm was used.

In RF algorithm, the mean decrease accuracy (MDA) index of each variable is determined when calculating the out-of-bag (OOB) error, which measures the importance of the variables by comparing how much OOB error of estimate value increases when excluding one variable and keeping others unchanged (Archer and Kimes, 2008; Verikas et al. 2011; Abdel-Rahman et al. 2013). Thus, the higher the MDA values of a variable, the greater its importance (Immitzer et al. 2012; Liu et al. 2017), we can thereby determine the most important variable. Additionally, compared with other algorithms, RF is more insensitive to multicollinearity, and its results are relatively robust to missing and unbalanced data, and it can well predict the effect of thousands of explanatory variables (Breiman 2001). Therefore, RF have been widely used in monitoring forest disturbance, especially for detecting

wood borer in pine forest (Abdel-Rahman et al. 2014; Lin et al. 2019; Iordache et al. 2020).

Currently, deep learning algorithms, such as convolutional neural network (CNN), have been showing its great potential in plant health monitoring (Yuan et al. 2017; Nagasubramanian et al. 2019; Wu et al. 2021). However, it still has some dependencies. Firstly, when the data is small, deep learning algorithms do not perform well. Furthermore, deep learning is like a black box, it does not reveal why it given the result, so it is lack of interpretability (Ling et al. 2018; Silaparasetty 2020). On the other side, with its rigorous calculations and great flexibility (Schmidhuber 2015; Hao et al. 2016), it could improve our classification accuracy. In our next study, deep learning algorithms will be employed on PWD diagnose using multi-temporal UAV-based hyperspectral data.

The possible application of Lidar

In this study, the classification model, predictions for Cab and WC, and the results of individual tree crown segmentation were obtained based on hyperspectral data alone. However, the results were not satisfactory, especially the tree crown segmentation (only delineated 121 from 218). Another potential method of data collection is Lidar (light detection and ranging). Lidar can directly, quickly, and accurately obtain three-dimensional geographic coordinates of objects (Vierling et al. 2008). Much progress has been made in the application of Lidar technology in the fields of geology, forestry and

Table 10 Classification confusion matrix of random forest (RF) classifier using vegetation indices, red edge parameters, moisture indices, and combined all variables based on UAV hyperspectral data

Stage	Vegetation Indices						Red Edge Parameters					
	H	E	M	S	Total	UA (%)	H	E	M	S	Total	UA (%)
H	25	9	1	0	35	71.43	28	8	1	0	37	75.68
E	10	16	2	0	28	57.14	9	15	3	0	27	55.56
M	3	1	22	4	30	73.33	1	2	22	3	28	78.57
S	1	1	4	22	28	78.57	1	2	3	23	29	79.31
Total	39	27	29	26	121		39	27	29	26	121	
PA (%)	64.10	59.26	75.86	84.62	OA (%)	70.25	71.79	55.56	75.86	88.46	OA (%)	72.73
					Kappa	0.60					Kappa	0.63
Stage	Moisture Indices						Combined					
	H	E	M	S	Total	UA (%)	H	E	M	S	Total	UA (%)
H	24	9	2	1	36	66.67	28	8	0	0	36	77.78
E	10	14	3	2	29	48.28	9	15	2	0	26	57.69
M	3	2	20	4	29	68.97	1	2	24	3	30	80.00
S	2	2	4	19	27	70.37	1	2	3	23	29	79.31
Total	39	27	29	26	121		39	27	29	26	121	
PA (%)	61.54	51.85	68.97	73.08	OA (%)	63.64	71.79	55.56	82.76	88.46	OA (%)	74.38
					kappa	0.51					Kappa	0.66

ecology, such as the establishment of digital elevation model (DEM), the extraction of forest structure parameters, and the inversion of forest ecosystem function parameters (e.g., Watt et al. 2014; Huang and Lian 2015; Saarela et al. 2020; Xie et al. 2020).

This makes Lidar a possible candidate to improve measurement accuracy. Although Lidar data failed to accurately reflect the biochemical condition of tree crowns (e.g., Liu et al. 2017; Shi et al. 2018), it can be used as measure auxiliary data that produces three-dimensional tree canopy structures (e.g., Shendryk et al. 2016). Thus, combining Lidar with hyperspectral data for individual
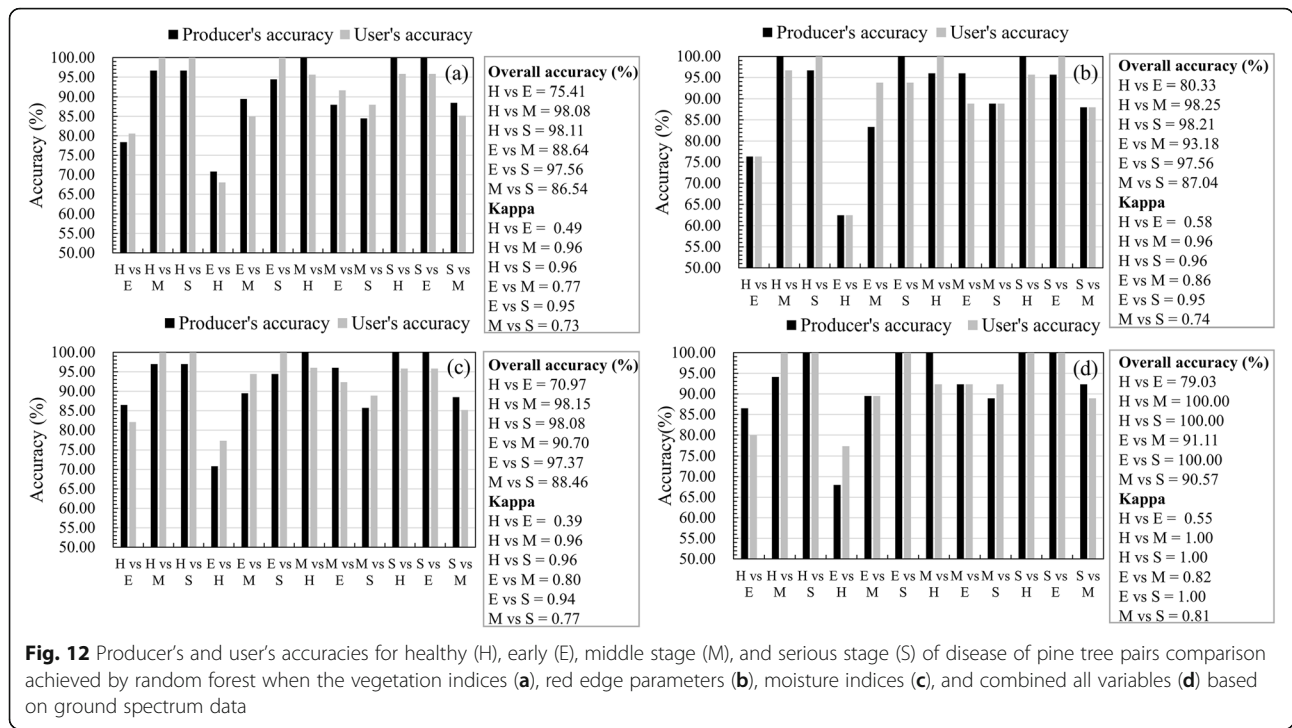
tree segmentation could improve accuracy (e.g., Junttila et al. 2019; Lin et al. 2019). Furthermore, crown structure and other tree structural information are likely to change throughout PWD infection. Therefore, variables based on the return intensity information from Lidar data might be useful in estimating the stage of PWD in pine trees, and it will be our next study.

Conclusion

In this paper, we compared the relatively accuracies of using ground-based data and UAV-based hyperspectral data in predicting the stage of PWD infection in pine
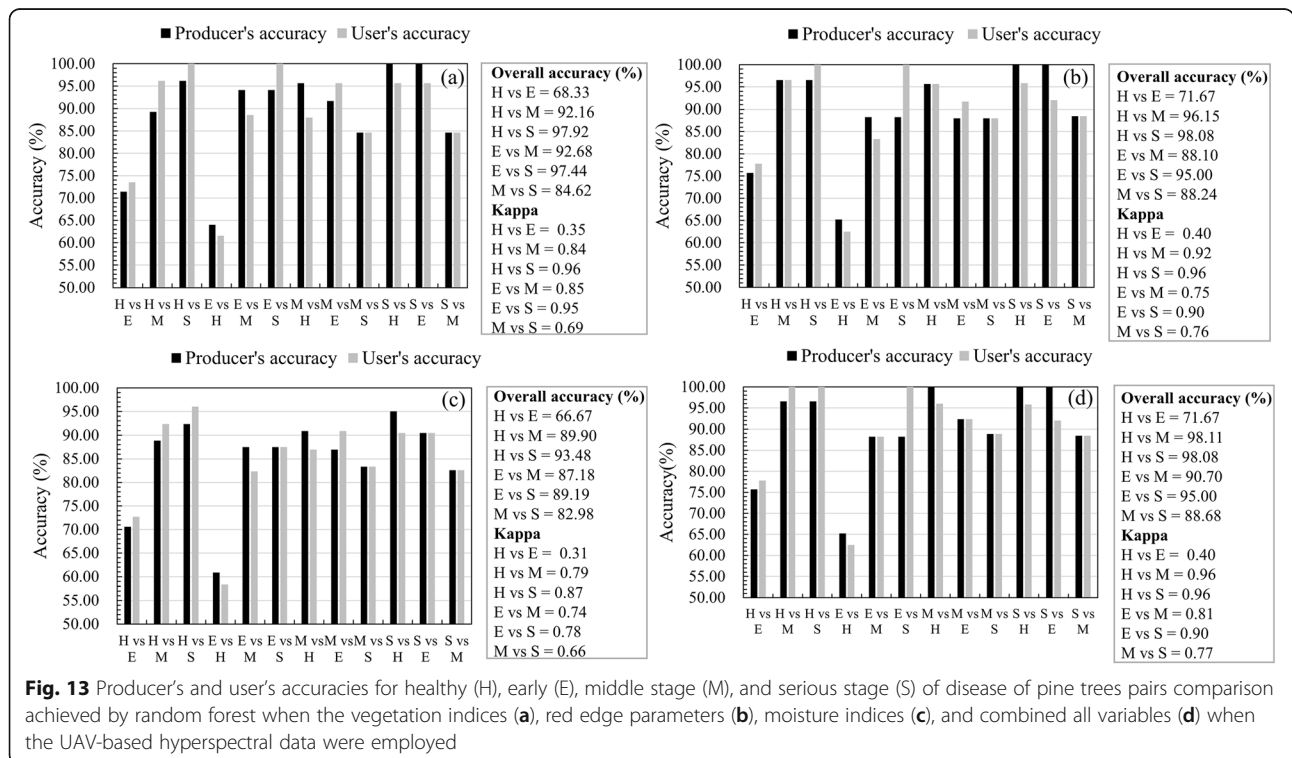
Table 11 Producer's accuracy (%) and user's accuracy (%) of the four stages using RF based on vegetation indices, red edge parameters, moisture indices, and combined all variables from ground and UAV-based data

Stage	Vegetation Indices		Red Edge Parameters		Moisture Indices		Combined	
	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy
Based on ground spectrum data								
H	74.36	80.56	82.05	82.05	74.36	74.36	82.05	80.00
E	62.96	60.71	62.96	73.91	55.56	57.69	62.96	70.83
M	75.86	78.57	82.76	80.00	82.76	80.00	82.76	80.00
S	88.46	79.31	88.46	79.31	84.62	84.62	92.31	88.89
UAV-based hyperspectral data								
H	64.10	71.43	71.79	75.68	61.54	66.67	71.79	77.78
E	59.26	57.14	55.56	55.56	51.85	48.28	55.56	57.69
M	75.86	73.33	75.86	78.57	68.97	68.97	82.76	80.00
S	84.62	78.57	88.46	79.31	73.08	70.37	88.46	79.31

Fig. 12 Producer's and user's accuracies for healthy (H), early (E), middle stage (M), and serious stage (S) of disease of pine tree pairs comparison achieved by random forest when the vegetation indices (**a**), red edge parameters (**b**), moisture indices (**c**), and combined all variables (**d**) based on ground spectrum data

trees. To do this, we selected VIs, REPs, MIs, and combining all variables as input parameters in a RF classification model. We found that combining all variables generally perform the best for estimating the stage of PWD infection of pine trees, and that REPs exhibit the highest accuracy in distinguishing between the healthy trees and trees in early stage of PWD infection. The classification accuracy of REPs based on UAV (airborne) data had slightly poorer performance in distinguishing trees at early stage of PWD and healthy trees (OA:

Fig. 13 Producer's and user's accuracies for healthy (H), early (E), middle stage (M), and serious stage (S) of disease of pine trees pairs comparison achieved by random forest when the vegetation indices (**a**), red edge parameters (**b**), moisture indices (**c**), and combined all variables (**d**) when the UAV-based hyperspectral data were employed

Yu *et al. Forest Ecosystems*   (2021) 8:44

Page 17 of 19

71.67%, Kappa: 0.40), but is still a feasible method. Therefore, UAV-based hyperspectral imaging is a promising candidate for measuring forest health. Relative to methods that use ground data, UAV-based hyperspectral imaging has the potential to substantially reduce labor and time costs. Future studies should aim to improve the accuracy of UAV-based data. One possible direction is the use of supplemental data acquisition practices such as UAV-based Lidar data to improve classification accuracy.

Availability of data and materials
The data are available upon a reasonable request to the Authors.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References
Abdel-Rahman EM, Ahmed FB, Ismail R (2013) Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. Int J Remote Sens 34(2):712–728. https://doi.org/10.1080/01431161.2012.713142

Abdel-Rahman EM, Mutanga O, Adam E, Ismail R (2014) Detecting *Sirex noctilio* grey-attacked and lightning-struck pine trees using airborne hyperspectral data, random forest and support vector machines classifiers. ISPRS J Photogrammd Remote Sens 88:48–59. https://doi.org/10.1016/j.isprsjprs.2013.11.013

Ahmed N, Atzberger C, Zewdie W (2020) Integration of remote sensing and bioclimatic data for prediction of invasive species distribution in data-poor regions: a review on challenges and opportunities. Environ Syst Res 9(1):32. https://doi.org/10.1186/s40068-020-00195-0

Archer KJ, Kimes RV (2008) Empirical characterization of random forest variable importance measures. Comput Stat Data Ann 52(4):2249–2260. https://doi.org/10.1016/j.csda.2007.08.015

Blackburn GA (1998) Quantifying chlorophylls and carotenoids at leaf and canopy scales: an evaluation of some hyperspectral approaches. Remote Sens Environ 66(3):273–285. https://doi.org/10.1016/S0034-4257(98)00059-5

Boochs F, Kupfer G, Dockter K, Kühbauch W (1990) Shape of the red edge as vitality indicator for plants. Int J Remote Sens 11(10):1741–1753. https://doi.org/10.1080/01431169008955127

Breiman L (2001) Random forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/a:1010933404324

Carter GA, Miller RL (1994) Early detection of plant stress by digital imaging within narrow stress-sensitive wavebands. Remote Sens Environ 50(3):295–302. https://doi.org/10.1016/0034-4257(94)90079-5

Chen JM (1996) Evaluation of vegetation indices and a modified simple ratio for boreal applications. Can J Remote Sens 22(3):229–242. https://doi.org/10.1080/07038992.1996.10855178

Chen YH (2005) Effects of pine wood nematode (PWN) infection on water regime and metabolism of related to hosts. Acta Phytopathol Sin 35(3):201–207. https://doi.org/10.13926/j.cnki.apps.2005.03.003 (in Chinese)

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens Environ 37(1):35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Curran PJ, Dungan JL, Gholz HL (1990) Exploring the relationship between reflectance red edge and chlorophyll content in slash pine. Tree Physiol 7(1-2-3-4):33–48. https://doi.org/10.1093/treephys/7.1-2-3-4.33

Dawson TP, Curran PJ (1998) A new technique for interpolating the reflectance red edge position. Int J Remote Sens 19(11):2133–2139. https://doi.org/10.1080/014311698214910

De Klerk HM, Buchanan G (2017) Remote sensing training in African conservation. Remote Sens Ecol Conserv 3(1):7–20. https://doi.org/10.1002/rse2.36

Douda O, Zouhar M, Maňasová M, Dlouhý M, Lišková J, Ryšánek P (2015) Hydrogen cyanide for treating wood against pine wood nematode (*Bursaphelenchus xylophilus*): results of a model study. J Wood Sci 61(2):204–210. https://doi.org/10.1007/s10086-014-1452-9

Du HQ, Ge HL, Fan YW, Jin W, Li J (2009) Application of fractal theory in hyperspectral detecting the early stage of pine wood nematode disease (*Bursaphelenchus xylophilus*) of *Pinus massoniana* with Hyperspectrum. Sci Silv Sin 45(6):68–76. https://doi.org/10.3321/j.issn:1001-7488.2009.06.012 (in Chinese)

Franklin SE, Wulder MA, Skakun RS, Carroll AL (2003) Mountain pine beetle red-attack forest damage classification using stratified Landsat TM data in British Columbia, Canada. Photogr Eng Remote Sens 69:283–288. https://doi.org/10.14358/pers.69.3.283

Gao B-C (1996) NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens Environ 58(3):257–266. https://doi.org/10.1016/s0034-4257(96)00067-3

Hao X, Zhang G, Ma S (2016) Deep learning. Int J Semantic Comput 10(3):417–439. https://doi.org/10.1142/S1793351X16500045

Hardisky MA, Smart RM, Klemas V (1983) Seasonal spectral characteristics and aboveground biomass of the tidal marsh plant, *Spartina alterniflora*. Photogramm Eng Remote Sens 49:85–92

He KS, Bradley BA, Cord AF, Rocchini D, Tuanmu M-N, Schmidtlein S, Turner W, Wegmann M, Pettorelli N (2015) Will remote sensing shape the next generation of species distribution models? Remote Sens Ecol Conserv 1(1):4–18. https://doi.org/10.1002/rse2.7

Hicke JA, Logan J (2009) Mapping whitebark pine mortality caused by a mountain pine beetle outbreak with high spatial resolution satellite imagery. Int J Remote Sens 30(17):4427–4441. https://doi.org/10.1080/01431160802566439

Horler DNH, Barber J, Barringer AR (1980) Effects of heavy metals on the absorbance and reflectance spectra of plants. Int J Remote Sens 1(2):121–136. https://doi.org/10.1080/01431160108559256

Huang BH (2020) Monitoring *Bursaphelenchus xylophilus* with multispectrum camera in UAV. Guangxi Forest Sci 49(3):380–384. https://doi.org/10.19692/j.cnki.gfs.2020.03.012 (in Chinese)

Huang H, Lian J (2015) A 3D approach to reconstruct continuous optical images using lidar and MODIS. Forest Ecosyst 2(1):20. https://doi.org/10.1186/s40663-015-0044-5

Huang HH, Ma XH, Huang HY, Zhou YF, Zhang W, Huang YH (2018) A preliminary study on monitoring of dead pine trees caused by pine wilt disease with fixed-wing unmanned aerial vehicle. J Environ Entomol 40(2):306–313. https://doi.org/10.3969/j.issn.1674-0858.2018.02.9 (in Chinese)

Huang MX, Gong JH, Li S, Zhang B, Hao QT (2012) Study on pine wilt disease hyper-spectral time series and sensitive features. Remote Sens Technol Appl 27(6):954–960. https://doi.org/10.11873/j.issn.1004-0323.2012.6.954 (in Chinese)

Hui J (2018) Damage and control measures of pine wilt disease. China South Agric Machine 49(4):100. https://doi.org/10.3969/j.issn.1672-3872.2018.04.087 (in Chinese)

Hunt ER, Rock BN (1989) Detection of changes in leaf water content using near- and middle-infrared reflectances. Remote Sens Environ 30(1):43–54. https://doi.org/10.1016/0034-4257(89)90046-1

Hyun MW, Kim JH, Suh DY, Lee SK, Kim SH (2007) Fungi isolated from pine wood nematode, its vector Japanese pine sawyer, and the nematode-infected Japanese black pine wood in Korea. Mycobiology 35(3):159–161. https://doi.org/10.4489/MYCO.2007.35.3.159

Immitzer M, Atzberger C, Koukal T (2012) Tree species classification with random forest using very high spatial resolution 8-band worldview-2 satellite data. Remote Sens 4(9):2661–2693. https://doi.org/10.3390/rs4092661

Inoue Y, Sakaiya E, Zhu Y, Takahashi W (2012) Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. Remote Sens Environ 126:210–221. https://doi.org/10.1016/j.rse.2012.08.026

Iordache M-D, Mantas V, Baltazar E, Pauly K, Lewyckyj N (2020) A machine learning approach to detecting pine wilt disease using airborne spectral imagery. Remote Sens 12(14):2280. https://doi.org/10.3390/rs12142280

Jones HG, Vaughan RA (2010) Remote sensing of vegetation. Oxford University Press, New York, pp 169–171

Jung KY, Park JK (2019) Analysis of vegetation infection information using unmanned aerial vehicle with optical sensor. Sensor Material 31:3319–3326. https://doi.org/10.18494/sam.2019.2465

Junttila S, Holopainen M, Vastaranta M, Lyytikäinen-Saarenmaa P, Kaartinen H, Hyyppä J, Hyyppä H (2019) The potential of dual-wavelength terrestrial lidar in early detection of *Ips typographus* (L.) infestation – leaf water content as a proxy. Remote Sens Environ 231:111264. https://doi.org/10.1016/j.rse.2019.111264

Kim S-R, Lee W-K, Lim C-H, Kim H, Kafatos MC, Lee S-H, Lee S-S (2018) Hyperspectral analysis of pine wilt disease to determine an optimal detection index. Forests 9(3):115. https://doi.org/10.3390/f9030115

Kuai CL (2012) Occurrence and control of pine wilt disease. Modern Agric Sci Technol 18:123–123. https://doi.org/10.3969/j.issn.1007-5739.2012.18.079 (in Chinese)

Li H, Xu HH, Zheng HY, Chen XY (2020) Monitoring technology of pine wilt disease based on UAV remote sensing image. J Chin Agric Mechan 41(9): 170–175. https://doi.org/10.13733/j.jcam.issn.2095-5553.2020.09.027 (in Chinese)

Li H, Zhou GY, Liu JN, Zhang HY (2011) Study on pine wilt disease and its control situation. Appl Mechd Mater 55–57:567–572. https://doi.org/10.4028/www.scientific.net/amm.55-57.567

Li ZX, Zhang XL, Zhang SC, Yan XS (2004) Application of RS and GIS in monitoring forest disease and insect pests. Hebei J For Orchard Res 19(4): 377–380. https://doi.org/10.3969/j.issn.1007-4961.2004.04.017 (in Chinese)

Lin Q, Huang H, Wang J, Huang K, Liu Y (2019) Detection of pine shoot beetle (PSB) stress on pine forests at individual tree level using UAV-based hyperspectral imagery and lidar. Remote Sens 11(21):2540. https://doi.org/10.3390/rs11212540

Lin XZ (2015) Review on damage and control measures of pine wilt disease. East China Forest Manag 29(3):28–30. https://doi.org/10.3969/j.issn.1004-7743.2015.03.008 (in Chinese)

Ling G, Lei G, Elmadany NEI, Liang C (2018) Statistical machine learning vs deep learning in information fusion: competition or collaboration? In: 2018 IEEE conference on multimedia information processing and retrieval. MIPR, Miami, pp 251–256. https://doi.org/10.1109/MIPR.2018.00059

Liu L, Coops NC, Aven NW, Pang Y (2017) Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. Remote Sens Environ 200:170–182. https://doi.org/10.1016/j.rse.2017.08.010

Liu W, Chang QR, Guo M, Xing DX, Yuan YS (2010) Monitoring of leaf nitrogen content in summer corn with first derivative of spectrum based on modified red edge. J Northwest A F Univ 38(4):91–98. https://doi.org/10.13207/j.cnki.jnwafu.2010.04.018 (in Chinese)

Ma J, Liu W, Zhang XL (2011) Remote sensing research on early monitoring and prediction of pine wilt disease. Forest Invent Plan 36(5):75–80. https://doi.org/10.3969/j.issn.1671-3168.2011.05.018 (in Chinese)

Mamiya Y (1988) History of pine wilt disease in Japan. J Nematol 20(2):219–226

Meddens AJH, Hicke JA, Vierling LA (2011) Evaluating the potential of multispectral imagery to map multiple stages of tree mortality. Remote Sens Environ 115(7):1632–1642. https://doi.org/10.1016/j.rse.2011.02.018

Mullen KE (2016) Early detection of mountain pine beetle damage in ponderosa pine forests of the black hills using hyperspectral and WorldView-2 data. Minnesota State University, Mankato, pp 58–60

Mutanga O, Adam E, Cho MA (2012) High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. Int J Appl Earth Observ Geoinform 18:399–406. https://doi.org/10.1016/j.jag.2012.03.012

Nagasubramanian K, Jones S, Singh AK, Sarkar S, Singh A, Ganapathysubramanian B (2019) Plant disease identification using explainable 3D deep learning on hyperspectral images. Plant Methods 15(1):98. https://doi.org/10.1186/s13007-019-0479-8

Näsi R, Honkavaara E, Lyytikäinen-Saarenmaa P, Blomqvist M, Litkey P, Hakala T, Viljanen N, Kantola T, Tanhuanpää T, Holopainen M (2015) Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level. Remote Sens 7(11):15467–15493. https://doi.org/10.3390/rs71115467

National Forestry Administration (2018) National Forestry Administration Announcement No.1 in 2018. http://www.forestry.gov.cn/main/3457/20180207/1074330.html. Accessed 16 Oct 2020

Pan CS (2011) Development of studies on pinewood nematodes diseases. J Xiamen Univ 50(2):476–483 (in Chinese)

Pan J, Ju YW, Wang XT, Zhang H (2014) Detection of *Bursaphelenchus xylophilus* infection in *Pinus massoniana* from hyperspectral data. Nematology 16(10): 1197–1207. https://doi.org/10.1163/15685411-00002846

Pan L, Li YX, Liu ZK, Meng FL, Chen J, Zhang XY (2019) Isolation and identification of pine wood nematode in *Pinus koraiensis* in Fengcheng, Liaoning Province. Forest Pest Dis 38(1):1–4. https://doi.org/10.19688/j.cnki.issn1671-0886.20180021 (in Chinese)

Penuelas J, Pinol J, Ogaya R, Filella I (1997) Estimation of plant water concentration by the reflectance water index WI (R900/R970). Int J Remote Sens 18(13):2869–2875. https://doi.org/10.1080/014311697217396

Pontius J, Martin M, Plourde L, Hallett R (2008) Ash decline assessment in emerald ash borer-infested regions: a test of tree-level, hyperspectral technologies. Remote Sens Environ 112(5):2665–2676. https://doi.org/10.1016/j.rse.2007.12.011

Prasad AM, Iverson LR, Liaw A (2006) Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9(2):181–199. https://doi.org/10.1007/s10021-005-0054-1

Richardson AJ, Wiegand CL (1977) Distinguishing vegetation from soil background information. Photogram Eng Remote Sens 43(12):1541–1552

Saarela S, Wästlund A, Holmström E, Mensah AA, Holm S, Nilsson M, Fridman J, Ståhl G (2020) Mapping aboveground biomass and its prediction uncertainty using LiDAR and field data, accounting for tree-level allometric and LiDAR model errors. Forest Ecosyst 7(1):43. https://doi.org/10.1186/s40663-020-00245-0

Santos CS, de Vasconcelos MW (2012) Identification of genes differentially expressed in *Pinus pinaster* and *Pinus pinea* after infection with the pine wood nematode. Eur J Plant Pathol 132(3):407–418. https://doi.org/10.1007/s10658-011-9886-z

Santoso H, Tani H, Wang X (2016) A simple method for detection and counting of oil palm trees using high-resolution multispectral satellite imagery. Int J Remote Sens 37(21):5122–5134. https://doi.org/10.1080/01431161.2016.1226527

Schlemmer M, Gitelson A, Schepers J, Ferguson R, Peng Y, Shanahan J, Rundquist D (2013) Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. Int J Appl Earth Observ Geoinform 25:47–54. https://doi.org/10.1016/j.jag.2013.04.003

Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Shen GR, Wang RC (2001) Review of the application of vegetation remote sensing. J Zhejiang Univ (Agric Life Sci) 27(6):682–690. https://doi.org/10.3321/j.issn.1008-9209.2001.06.025 (in Chinese)

Shendryk I, Broich M, Tulbure MG, McGrath A, Keith D, Alexandrov SV (2016) Mapping individual tree health using full-waveform airborne laser scans and imaging spectroscopy: a case study for a floodplain eucalypt forest. Remote Sens Environ 187:202–217. https://doi.org/10.1016/j.rse.2016.10.014

Shi Y, Skidmore AK, Wang T, Holzwarth S, Heiden U, Pinnel N, Zhu X, Heurich M (2018) Tree species classification using plant functional traits from LiDAR and hyperspectral data. Int J Appl Earth Observ Geoinform 73:207–219. https://doi.org/10.1016/j.jag.2018.06.018

Shin S-C (2008) Pine wilt disease in Korea. In: Zhao BG, Futai K, Sutherland JR, Takeuchi Y (eds) Pine wilt disease. Springer, Tokyo. https://doi.org/10.1007/978-4-431-75655-2_5

Silaparasetty N (2020) Machine learning vs. deep learning. In: Silaparasetty N (ed) Machine learning concepts with Python and the Jupyter notebook environment. Using Tensorflow 2.0. Apress, Berkeley, pp 57–65. https://doi.org/10.1007/978-1-4842-5967-2_4

Song QL, Xiang R, Qing L, Yuan L, Tian H, Zhu PB, Liu W, Yang W, Qu YX, Zhou JW (2018) Relationship between spectral reflectance characteristic parameters and water content of pine leaves. Sci Technol Innov 03:26–28. https://doi.org/10.15913/j.cnki.kjycx.2018.03.026 (in Chinese)

Syifa M, Park S-J, Lee C-W (2020) Detection of the pine wilt disease tree candidates for drone remote sensing using artificial intelligence techniques. Engineering 6(8):919–926. https://doi.org/10.1016/j.eng.2020.07.001

Tang L, Shao G (2015) Drone remote sensing for forestry research and practices. J For Res 26(4):791–797. https://doi.org/10.1007/s11676-015-0088-y

Tao H, Li C, Zhao D, Deng S, Hu H, Xu X, Jing W (2020) Deep learning-based dead pine trees detection from unmanned aerial vehicle images. Int J Remote Sens 41(21):8238–8255. https://doi.org/10.1080/01431161.2020.1766145

Umebayashi T, Yamada T, Fukuhara K, Endo R, Kusumoto D, Fukuda K (2017) In situ observation of pinewood nematode in wood. Eur J Plant Pathol 147(2):463–467. https://doi.org/10.1007/s10658-016-1013-8

Verikas A, Gelzinis A, Bacauskiene M (2011) Mining data with random forests: a survey and results of new tests. Pattern Recogn 44(2):330–349. https://doi.org/10.1016/j.patcog.2010.08.011

Vicente C, Espada M, Vieira P, Mota M (2012) Erratum to: pine wilt disease: a threat to European forestry. Eur J Plant Pathol 133(2):497. https://doi.org/10.1007/s10658-012-9979-3

Vierling KT, Vierling LA, Gould WA, Martinuzzi S, Clawges RM (2008) Lidar: shedding new light on habitat characterization and modeling. Front Ecol Environ 6(2):90–98. https://doi.org/10.1890/070001

Wang Z, Zhang X, An S (2007) Spectral characteristics analysis of *Pinus Massoniana* suffered by *Bursaphelenchus Xylophilus*. Remote Sens Technol Appl 22(3):367–370. https://doi.org/10.3969/j.issn.1004-0323.2007.03.012 (in Chinese)

Waske B, Benediktsson JA, Árnason K, Sveinsson JR (2009) Mapping of hyperspectral AVIRIS data using machine-learning algorithms. Can J Remote Sens 35(Sup1):S106–S116. https://doi.org/10.5589/m09-018

Watt MS, Meredith A, Watt P, Gunn A (2014) The influence of LiDAR pulse density on the precision of inventory metrics in young unthinned Douglas-fir stands during initial and subsequent LiDAR acquisitions. NZ J Forestry Sci 44(1):18. https://doi.org/10.1186/s40490-014-0018-3

White JC, Coops NC, Hilker T, Wulder MA, Carroll AL (2007) Detecting mountain pine beetle red attack damage with EO-1 Hyperion moisture indices. Int J Remote Sens 28(10):2111–2121. https://doi.org/10.1080/01431160600944028

White JC, Wulder MA, Brooks D, Reich R, Wheate RD (2005) Detection of red attack stage mountain pine beetle infestation with high spatial resolution satellite imagery. Remote Sens Environ 96(3–4):340–351. https://doi.org/10.1016/j.rse.2005.03.007

Wu B, Liang A, Zhang H, Zhu T, Zou Z, Yang D, Tang W, Li J, Su J (2021) Application of conventional UAV-based high-throughput object detection to the early diagnosis of pine wilt disease by deep learning. Forest Ecol Manag 486:118986. https://doi.org/10.1016/j.foreco.2021.118986

Wu CY, Niu Z (2008) Sensitivity study of photochemical reflectance index to leaf biochemical components. J Univ Chin Acad Sci 3:346–354 (in Chinese)

Xiang R, Yuan L, Qin L, Song Q, Zhang J, Qu Y, Zhou J (2018) Correlation analysis between spectral characteristic parameters and chlorophyll content of pine needles. Sci Technol Innov 35:5–7 (in Chinese)

Xie Q, Huang W, Liang D, Chen P, Wu C, Yang G, Zhang J, Huang L, Zhang D (2014) Leaf area index estimation using vegetation indices derived from airborne hyperspectral images in winter wheat. IEEE J-STARS 7(8):3586–3594. https://doi.org/10.1109/JSTARS.2014.2342291

Xie Y, Zhang J, Chen X, Pang S, Zeng H, Shen Z (2020) Accuracy assessment and error analysis for diameter at breast height measurement of trees obtained using a novel backpack LiDAR system. Forest Ecosyst 7(1):33. https://doi.org/10.1186/s40663-020-00237-0

Xu HC, Luo YQ, Zhang Q (2012) Changes in water content, pigments and antioxidant enzyme activities in pine needles of *Pinus thunbergii* and *Pinus massoniana* affected by pine wood nematode. Sci Silv Sin 11:140–143 (in Chinese)

Xu HC, Luo YQ, Zhang TT, Shi YJ (2011) Changes of reflectance spectra of pine needles in different stage after being infected by pine wood nematode.

Spectrosc Spectr Anal 31(5):1352–1356. https://doi.org/10.3964/j.issn.1000-0593(2011)05-1352-05 (in Chinese)

Yang BJ (2002) Advance in research of pathogenetic mechanism of pine wood nematode. Forest Pest Disease 1:27–31 (in Chinese)

Yao FQ, Zhang ZH, Yang RY, Sun JW, Cui SF (2009) Hyperspectral models for estimating vegetation chlorophyll content based on red edge parameter. Trans Chin Soc Agric Eng 25(Sup2):123–129 (in Chinese)

Ye JR (2019) Epidemic status of pine wilt disease in China and its prevention and control techniques and counter measures. Sci Silv Sin 55(9):1–10 (in Chinese)

Yu HY, Wu H (2018) Discovery of new host plants and new vector insects of *Bursaphelenchus xylophilus* in Liaoning Province. Forest Pest Dis 37(5):61 (in Chinese)

Yu HY, Wu H, Zhang XD, Wang LM, Zhang XF, Song YS (2019) Preliminary study on *Larix* spp. infected by *Bursaphelenchus xylophilus* in natural environment. Forest Pest Dis 38(4):7–10. https://doi.org/10.19688/j.cnki.issn1671-0886.20180024 (in Chinese)

Yuan H, Yang G, Li C, Wang Y, Liu J, Yu H, Feng H, Xu B, Zhao X, Yang X (2017) Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: analysis of RF, ANN, and SVM regression models. Remote Sens 9(4):309. https://doi.org/10.3390/rs9040309

Yuan J, Wang DL, Li R (2013) Remote sensing image segmentation by combining spectral and texture features. IEEE T Geosci Remote 52(1):16–24. https://doi.org/10.1109/TGRS.2012.2234755

Zhan Z, Yu L, Li Z, Rem L, Gao B, Wang L, Luo Y (2020) Combining GF-2 and Sentinel-2 images to detect tree mortality caused by red turpentine beetle during the early outbreak stage in North China. Forests 11(2):172. https://doi.org/10.3390/f11020172

Zhang N, Zhang X, Yang G, Zhu C, Huo L, Feng H (2018) Assessment of defoliation during the *Dendrolimus tabulaeformis Tsai et Liu* disaster outbreak using UAV-based hyperspectral images. Remote Sens Environ 217:323–339. https://doi.org/10.1016/j.rse.2018.08.024